

Marquette University

e-Publications@Marquette

Electrical and Computer Engineering Faculty
Research and Publications

Electrical and Computer Engineering,
Department of

2-2006

Detection Efficiencies and Generalized Breakdown Probabilities for Nanosecond-Gated Near Infrared Single-Photon Avalanche Photodiodes

David A. Ramirez

University of Concepción

Majeed M. Hayat

Marquette University, majeed.hayat@marquette.edu

Gauri Karve

University of Texas - Austin

Joe C. Campbell

University of Texas

Sergio N. Torres

University of Concepcion

See next page for additional authors

Follow this and additional works at: https://epublications.marquette.edu/electric_fac



Part of the [Computer Engineering Commons](#), and the [Electrical and Computer Engineering Commons](#)

Recommended Citation

Ramirez, David A.; Hayat, Majeed M.; Karve, Gauri; Campbell, Joe C.; Torres, Sergio N.; Saleh, Bahaa E.A.; and Teich, Malvin Carl, "Detection Efficiencies and Generalized Breakdown Probabilities for Nanosecond-Gated Near Infrared Single-Photon Avalanche Photodiodes" (2006). *Electrical and Computer Engineering Faculty Research and Publications*. 626.

https://epublications.marquette.edu/electric_fac/626

Authors

David A. Ramirez, Majeed M. Hayat, Gauri Karve, Joe C. Campbell, Sergio N. Torres, Bahaa E.A. Saleh, and Malvin Carl Teich

Marquette University

e-Publications@Marquette

Electrical and Computer Engineering Faculty Research and Publications/College of Engineering

This paper is NOT THE PUBLISHED VERSION; but the author's final, peer-reviewed manuscript. The published version may be accessed by following the link in the citation below.

IEEE Journal of Quantum Electronics, Vol. 42, No. 2 (February 2006): 137-145. [DOI](#). This article is © Institute of Electrical and Electronic Engineers (IEEE) and permission has been granted for this version to appear in [e-Publications@Marquette](#). Institute of Electrical and Electronic Engineers (IEEE) does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Institute of Electrical and Electronic Engineers (IEEE).

Detection Efficiencies and Generalized Breakdown Probabilities for Nanosecond-Gated Near Infrared Single-Photon Avalanche Photodiodes

David A. Ramirez

Department of Electrical Engineering, University of Concepcion, Chile

Majeed M. Hayat

Department of Electrical and Computer Engineering at the University of New Mexico, Albuquerque

Gauri Karve

University of Texas at Austin

Joe C. Campbell

University of Texas

Sergio N. Torres

Electrical Engineering Department, University of Concepcion

Bahaa E. A. Saleh

Department of Electrical and Computer Engineering at Boston University, Boston, MA

Malvin C. Teich

Departments of Electrical and Computer Engineering, Physics, and Biomedical Engineering, Boston University, Boston, MA

Abstract

A rigorous model is developed for determining single-photon quantum efficiency (SPQE) of single-photon avalanche photodiodes (SPADs) with simple or heterojunction multiplication regions. The analysis assumes nanosecond gated-mode operation of the SPADs and that band-to-band tunneling of carriers is the dominant source of dark current in the multiplication region. The model is then utilized to optimize the SPQE as a function of the applied voltage, for a given operating temperature and multiplication-region structure and material. The model can be applied to SPADs with In/sub 0.52/Al/sub 0.48/As or InP multiplication regions as well as In/sub 0.52/Al/sub 0.48/As--InP heterojunction multiplication regions for wavelengths of 1.3 and 1.55 μm . The predictions show that the SPQE generally decreases with decreasing the multiplication-region thickness. Moreover, an InP multiplication region requires a lower breakdown electric field (and, hence, offers a higher SPQE) than that required by an In/sub 0.52/Al/sub 0.48/As layer of the same width. The model also shows that the fractional width of the In/sub 0.52/Al/sub 0.48/As layer in an In/sub 0.52/Al/sub 0.48/As--InP heterojunction multiplication region can be optimized to attain a maximum SPQE that is greater than that offered by an InP multiplication region. This effect becomes more pronounced in thin multiplication regions as a result of the increased significance of dead space.

SECTION I. Introduction

SINGLE-PHOTON avalanche photodiodes (SPADs), also known as Geiger-mode avalanche photodiodes (GM-APDs), are important devices for visible-light to near-infrared applications that demand high single-photon detection efficiencies. These applications include satellite laser ranging [1], deep-space laser communication [2], time-resolved photon counting [3], quantum key distribution [4], quantum imaging [5], and quantum cryptography [6], [7].

Silicon APDs have already shown very good performance in various applications in the 400–900-nm range. When operated in Geiger mode, silicon APDs can offer detection efficiencies around 60%, dark counts that are below 100 counts per second, and subnanosecond timing resolution [8], [9]. However, their performance is degraded drastically when they are operated in the range 1.06–1.55 μm , due to their low absorption at these wavelengths. For example, silicon SPADs are typically limited to below 20% quantum efficiency and 2% photon-detection efficiency at 1.06 μm [10]. At 1.3 μm , the quantum efficiency drops to $10^{-7}\%$ [11].

For applications in the range 1.3–1.55 μm (i.e., telecommunication wavelengths) devices with a narrower bandgap than silicon must be utilized, including germanium, silicon-germanium, and III–V compounds such as InGaAs–InP separate-absorption-multiplication (SAM) heterostructures [9], [12]–[13][14][15]. Although we have recently witnessed remarkable advances in low-noise APDs for linear-

mode operation in the 1.3–1.55- μm range, their performance in the Geiger mode remains unimpressive. Currently, one of the best reported photon counters is the EPM 239 InGaAs–InP APD, which is manufactured by the Epitaxx division of JDS-Uniphase [9]. This SPAD has a photon-detection efficiency of 10% at 1.55 μm . Clearly, at the present time there is a large gap between the performance of APDs in Geiger and linear modes for wavelengths that are of interest to telecommunications. This may be partially attributed to lack of known correspondence between the performance metrics and device optimization strategies for Geiger and linear modes of operation.

To date, only a heuristic connection exists between the performance metrics and criteria for linear and Geiger modes of APD operation. The performance of single-photon detection has heretofore not been so quantified. Linear-mode performance measures, such as the excess-noise factor and the gain-bandwidth product, are not suitable for characterizing breakdown characteristics nor the single-photon detection capability. In this paper, we present a general mathematical treatment of performance assessment and optimization of APDs for Geiger-mode operation. This establishes how modern APD structures, which are known to yield a performance advantage in linear-mode operation, perform in a Geiger-mode setting.

It has been well established that APDs with thin multiplication regions exhibit a significant reduction in the excess noise factor. This trend has been primarily attributable to the dead-space effect [16]–[17][18][19]. New structures have also been proposed and shown to exhibit even further reduction in the excess noise factor. For example, recent analytical studies, Monte Carlo simulations, and experimental studies have shown that the use of impact-ionization engineered (I^2E) multiplication regions can significantly reduce the excess noise factor in avalanche photodiodes [20], [21]. In fact, the lowest excess noise factor reported to date at 1.55 μm has been achieved with an (I^2E) APD [22].

The first step in understanding the performance of the aforementioned APDs in Geiger-mode operation is to assess their breakdown characteristics, *viz.*, the breakdown probability as a function of the applied reverse bias. Typically, the breakdown probability increases from zero to unity as the applied-reverse voltage is raised beyond a threshold called the breakdown voltage, V_{BR} . In general, the steeper the zero-to-one transition in the breakdown probability, as a function of the excess bias, the better are the breakdown characteristics. In simple multiplication regions, the breakdown characteristics improve as the width of the multiplication region increases [23], [24]. On the other hand, it has been shown that thin heterojunction multiplication regions can achieve the same breakdown characteristic as thicker, simple multiplication regions ($> 500\text{nm}$) [23]. However, an improvement in the breakdown characteristics does not necessarily imply an improvement in Geiger-mode performance, which is primarily characterized by the single-photon quantum efficiency (SPQE) [9], [25].

Aside from the breakdown characteristics, there are a number of other key factors, such as the magnitude of electric field and the dark-carrier generation in the multiplication region and the absorber, that can dominate the SPQE as the width of the multiplication region and structure is changed. In particular, an important factor that is included in the SPQE is the dark-count probability, P_d . It reflects the false counts due to triggers caused by dark carriers. To date, the net effects of changing the multiplication-region width and structure on the SPQE and P_d have not been systematically studied.

In this paper we develop a rigorous model for assessing the dependence of the SPQE on the SPAD's multiplication-region structure and width. This approach provides a valuable analytical tool for exploring and optimizing novel APD structures for single-photon detection and counting for SPADs operated in nanosecond-gated mode. This paper addresses the optimization of SPADs as a function of the operating voltage, and structure, as well as the width of the multiplication region, under the assumption that band-to-band tunneling of carriers is the dominant source of dark current in the multiplication region.

SECTION II. Single-Photon Quantum Efficiency (SPQE)

In this section we review the SPQE performance metric and discuss a generalized technique for calculating the breakdown probability in SPADs with either simple or heterojunction multiplication regions. We assume gated-mode operation where the detection time is the gate-on time. In order to avoid a high dark-count rate we assume short gate-on times. In this sense, it should be pointed out that our definition of SPQE is valid for short gate-on times. The gate-on time considered is 2 ns.

The SPQE is defined as the probability that a photon triggers an avalanche breakdown and no dark carrier triggers a breakdown given that an optical pulse is present and at least one photon impinges on the SPAD. Mathematically, it is defined as

$$\text{SPQE} = \frac{(1-P_d)P_{\text{opt}}}{p_o} \quad (1)$$

where P_d is the dark-count probability defined as the probability that at least one dark carrier successfully triggers an avalanche breakdown [25], P_{opt} is the probability that at least one photon triggers an avalanche breakdown, and p_o is the probability that at least one photon impinges on the SPAD during the detection time. By assuming Poisson statistics for the dark carriers, P_d can be expressed as [25]

$$P_d = 1 - e^{-Q_d N_d} \quad (2)$$

where Q_d is the probability that a dark carrier triggers an avalanche breakdown and N_d is the average number of dark carriers generated in the multiplication region during the detection time.

Similarly, by assuming Poisson photon statistics, we obtain

$$P_{\text{opt}} = 1 - e^{-\eta P_a N_o} \quad (3)$$

where N_o is the average number of photons per pulse, P_a is the probability of an avalanche breakdown caused by a photo-carrier that is injected into the multiplication region, and η is the quantum efficiency of the SPAD, which is the probability that an impinging photon is absorbed and converted into an electron-hole pair. Also, the probability that at least one photon is present in the optical pulse is given by $p_o = 1 - e^{-N_o}$.

In general, our definition of SPQE considers the presence of a pulse (and not just a single photon), which may contain more than one photon. However, one may also be interested in a definition of the SPQE for which the presence of a single photon is assumed in the detection time. However, the fact

that the average number of photons N_o is assumed 0.1 in all of our calculations makes the single-photon-based and pulse-based definitions of the SPQE approximately equivalent. Indeed a simple calculation shows that when $N_o \ll 1$, our pulse-based definition of the SPQE can be approximated by $\text{SPQE} \approx (1 - P_a)\eta P_a$, which is indeed the definition of the SPQE given that one and only one photon is present in the detection interval.

Note that the photon-triggered and dark-carrier-triggered breakdown probabilities P_a and Q_a , respectively, are different in general. This is because the breakdown probability corresponding to a carrier is dependent on the location where the carrier is generated or injected in the multiplication region. In a separate-absorption-multiplication (SAM) SPAD, photo-generated carriers are assumed to be injected into the multiplication region while a dark carrier can be generated at a random location inside the multiplication region. A dark carrier generated in the absorber, however, should be treated as having the same breakdown probability as that corresponding to a photo-generated carrier. We defer the details of calculating the breakdown probabilities P_a and Q_a to Section III-A.

SECTION III. Model for Breakdown Probability and Dark Current

A. Breakdown Probability

We define two types of breakdown probabilities, depending on the location of the carrier that triggers the avalanche breakdown. As described earlier, the *injected-carrier* breakdown probability, P_a , which is used in (3), reflects the breakdown probability caused by a carrier that is injected into the multiplication region. For example, in a SAM structure, P_a corresponds to breakdown triggered by carriers that are optically generated in the absorption layer as well as any dark carrier generated therein. On the other hand, the *distributed-carrier* breakdown probability, Q_a , represents the breakdown probability due to a dark carrier randomly generated *inside* the multiplication region. Let us assume that dark carriers are generated according to certain spatial distribution in the multiplication region. Then, since the breakdown probability due to any carrier is dependent on the birthplace of the carrier, Q_a is the breakdown probability averaged over all possible locations of the random birthplaces of the dark carriers in the multiplication region [23]. Interestingly, this distributed-type behavior has been examined earlier by Hakim et al. in the context of excess-noise-factor characterization in linear-mode operation [26]. We now introduce a formalism that can be used in conjunction with the recursive dead-space multiplication theory (DSMT) [27] to calculate P_a and Q_a .

Let us assume that an electron–hole pair is created at a random location in an arbitrary (single or multiple layer) multiplication region extending from $x = 0$ to $x = w$. Assume that electrons (holes) are transported in the positive (negative) x direction. Moreover, let $f(x)$ denote the probability density function of the birthplace of the electron–hole pair. We let $P_e(x)$ denote the probability that an electron born at location x triggers an avalanche breakdown. Similarly, $P_h(x)$ represents the probability that a hole born at location x triggers avalanche breakdown. Now, the probability that an electron–hole pair born at x collectively triggers breakdown is $P_b(x) = 1 - (1 - P_e(x))(1 - P_h(x)) = P_e(x) + P_h(x) - P_e(x)P_h(x)$. Thus, it follows that the average probability that an electron–hole pair, randomly generated in the interval $[0, w]$ according to the probability density function f , triggers breakdown is given by

$$P_b = \int_0^w (P_e(x) + P_h(x) - P_e(x)P_h(x))f(x)dx. (4)$$

To determine the quantities $P_e(x)$ and $P_h(x)$, we invoke the DSMT recursive technique reported in [27] [see (4) and (5) therein for the recursive equations that govern $P_e(x)$ and $P_h(x)$]. We can therefore recast (4) as

$$P_b = 1 - \int_0^w P_Z(x)P_Y(x)f(x)dx (5)$$

where $P_Z(x) = 1 - P_e(x)$ is the probability that the total electron and hole population resulting from a parent electron born at x is finite; similarly, $P_Y(x) = 1 - P_h(x)$ is the probability that the total electron and hole population resulting from a parent hole born at x is finite.

As a special case, if an electron is injected into the multiplication region, as the case of photo-generated or a dark electron transported from the absorption region, we take $f(x) = \delta(x)$, the Dirac delta function. In this case the injected-carrier breakdown probability P_a is simply $P_e(0)$, since $P_h(0) = 0$. Thus, $P_a = 1 - P_Z(0)$. A similar argument holds for holes entering the multiplication region at $x = w$: in this case we take $f(x) = \delta(x - w)$. Since $P_e(w) = 0$, we obtain $P_a = P_h(w) = 1 - P_Y(w)$ in the case of hole injection at $x = w$. On the other hand, if a carrier-pair is generated randomly and uniformly in $[0, w]$, then we take $f(x) = w^{-1}$ and obtain the distributed-breakdown probability $Q_a = 1 - w^{-1} \int_0^w P_Z(x)P_Y(x)dx$.

We next describe the model for dark-carrier generation in the multiplication regions developed by Karve et al. [28]. This model is reviewed briefly in this paper for completeness; it is described in more detail in [28].

B. Band-to-Band Tunneling Current in the Multiplication Region

There are several mechanisms that contribute to the dark-carrier concentration in a SPAD [25]. These include band-to-band tunneling, release of trapped carriers from recombination centers, which give rise to afterpulsing, and thermally generated carriers. The dominant mechanism in a specific SPAD will depend upon the structure and operating conditions such as the bias voltage, repetition rate, temperature, etc.

At high electric fields (i.e., in the multiplication region and for biases that are beyond breakdown), the dark-carrier generation due to band-to-band tunneling becomes very important for thin multiplication layers, as shown by Karve et al. [28]. (Moreover, as tunneling depends weakly on temperature, tunneling eventually becomes the major source of dark current as the temperature is reduced, even if it is not dominant at 300 K [8], [29]). Generally, tunneling current increases exponentially as the electric field increases [29]. In the case of $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, for example, the average number of dark carriers N_d generated in the multiplication region can be calculated according to [28]

$$N_d = C_1 V E_m e^{-C_2/E_m} (6)$$

where V is the voltage across the multiplication region and E_m is the maximum electric field in the multiplication region. The precise expressions for the constants C_1 and C_2 are described in detail in [28].

SECTION IV. Performance of SPADs With Simple $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ or InP Multiplication Regions

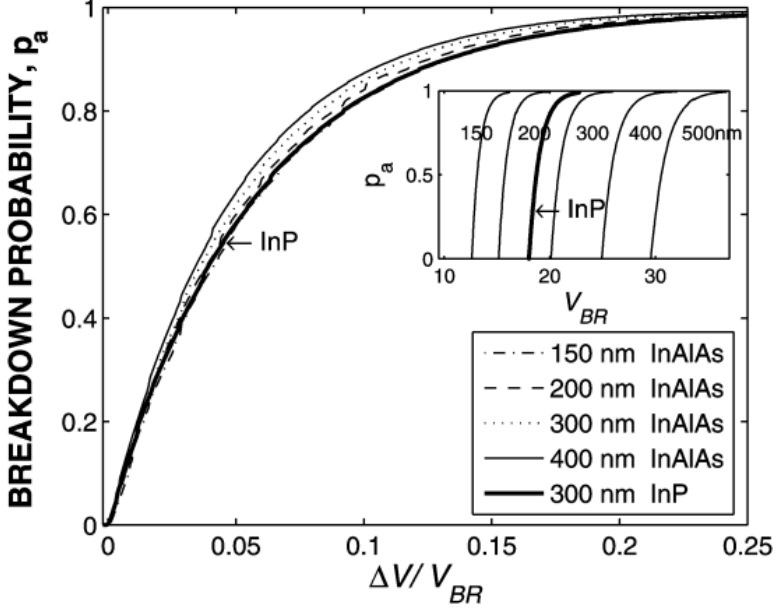


Fig. 1. Calculated injected-carrier breakdown probability for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ as a function of the normalized excess voltage for various multiplication-region widths. The breakdown probability for a 300-nm InP multiplication region is also shown for comparison. The Inset shows the breakdown probabilities as a function of the reverse-bias voltage.

In this section, we utilize the SPQE metric to predict the performance of SPADs with $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ and InP multiplication regions of various widths. In our calculations, we focus on the dark-carrier generation and breakdown characteristics only in the high electric field multiplication region. Throughout, we will assume that the SPAD is operated in a gated mode, whereby the SPAD is dc biased just below its breakdown voltage V_{BR} and then periodically pulse-biased above V_{BR} during the on-period [25]. We assume a 2-ns on-pulse with repetition rate of 500 kHz [28]. Moreover, the average number of photons per pulse is assumed as $N_o = 0.1$ photons. We choose such a low level of optical intensity so that the probability that two or more photons reaching the SPAD is negligible [14]. Finally, the quantum efficiency of the SPAD is assumed as $\eta = 0.5$ at $\lambda = 1.55 \mu\text{m}$ [30].

A. Dependence of the SPQE on the Multiplication-Region Width

We begin by presenting the dependence of the injected-carrier breakdown probability P_a on the multiplication-region width. The breakdown-probability characteristics are often assessed by plotting the breakdown probability as a function of the normalized excess voltage; a steep transition in the breakdown probability from 0 to 1 is always desirable. The normalized excess voltage is defined as $\Delta V/V_{\text{BR}}$, where $\Delta V = (V - V_{\text{BR}})$, V_{BR} is the breakdown voltage and V is the voltage across the multiplication region. It has been noted recently that as the width of the multiplication region is reduced, the breakdown-probability characteristics are degraded. In particular, Ng et al. [31] predicted that the breakdown-probability curves, as a function of the excess bias voltage, are steeper for thick multiplication regions than those for thin ones. Our calculations also confirm this trend for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ and InP , as shown in Fig. 1 for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ multiplication regions. This also confirms the already known fact that the breakdown voltage decreases as the width of the multiplication region

is reduced. Thus, for a fixed applied voltage, the breakdown probability of a thin layer is higher than that for a thick layer.

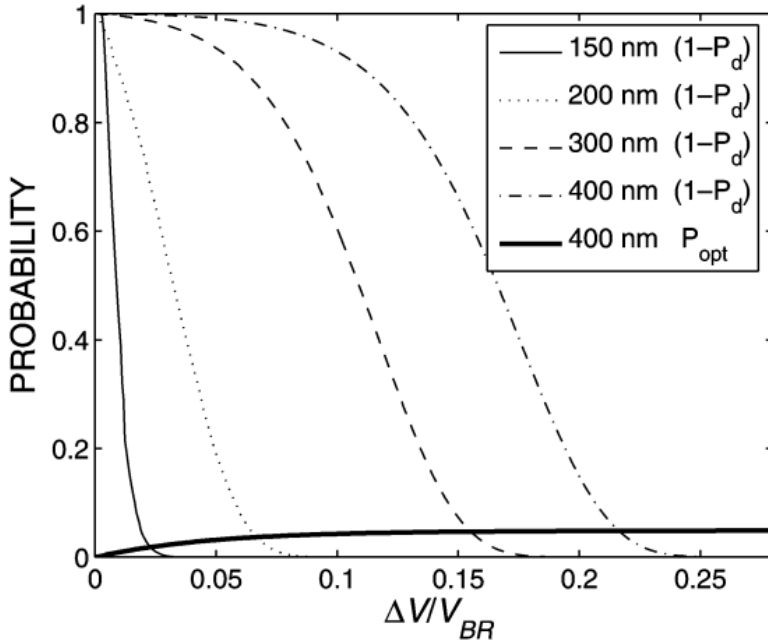


Fig. 2. Calculated probability $1 - P_d$ as a function of the normalized excess voltage for each multiplication-region width, w . The thick solid curve is P_{opt} as a function of the normalized excess voltage for $w = 400$ nm. P_{opt} for smaller widths are slightly below the 400-nm curve and are not shown in the figure.

As a result of the degradation in the breakdown probability characteristics, we would intuitively suspect that reducing the width of the multiplication region may also have a detrimental effect on the SPQE. However, no theoretical studies, to the best of our knowledge, have systematically shown the trend in the SPQE as a function of the width of the multiplication region. The subtlety in drawing any heuristic conclusions on the SPQE based on the breakdown-probability characteristics alone is that the average dark-count probability P_d , which is a key factor in SPQE, depends on the width in two ways. First, reducing the width results in an increase in the breakdown field, which, in turn, yields higher dark-carrier generation. On the other hand, due to the degradation in the breakdown characteristics as the width is reduced, the probability that a dark carrier triggers breakdown is also reduced. Since these two effects are at odds, it is not clear what the net effect on the SPQE due to reducing the multiplication-region width would be.

To rigorously address this question, we begin by examining Fig. 2, which depicts the behavior of the two terms $1 - P_d$ and P_{opt} in SPQE as a function of the normalized excess voltage for various $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ multiplication regions. (Recall that $1 - P_d$ is the probability that a dark carrier does not trigger an avalanche breakdown while P_{opt} is the probability that a photon triggers a breakdown.) Let us first examine the effect of changing w on $1 - P_d$. It can be noticed from Fig. 2 that increasing the multiplication-region width causes the $1 - P_d$ term to increase significantly. This desirable effect is essentially due to the reduction in the breakdown electric field in thick multiplication regions and it dominates the harmful effect caused by the increase in the breakdown probability as the width is increased. Our assumption that dark-carrier generation is dominated by tunneling may have

accentuated this effect due to the strong dependence of band-to-band tunneling current on the electric field. On the other hand, increasing the multiplication-region width has a relatively much less significant effect on P_{opt} compared to $1 - P_d$. Therefore, we would expect a net increase in the SPQE as the multiplication-region width increases, as shown in Fig. 3 for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ multiplication regions. Note that the theoretical limit of the SPQE is η , which is 0.5 in the examples considered.

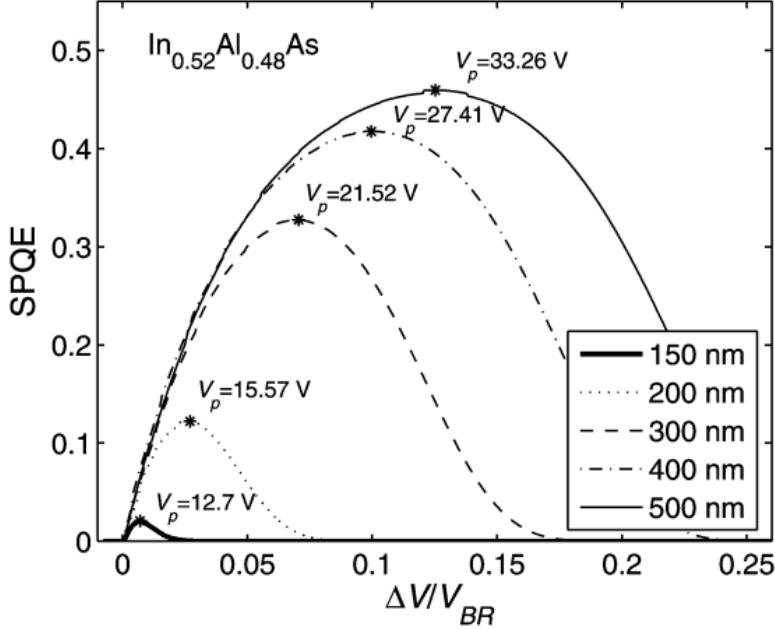


Fig. 3. Calculated SPQE as a function of the normalized excess voltage for each multiplication-region width. The reverse-bias voltage corresponding to the peak SPQE is also shown for each width.

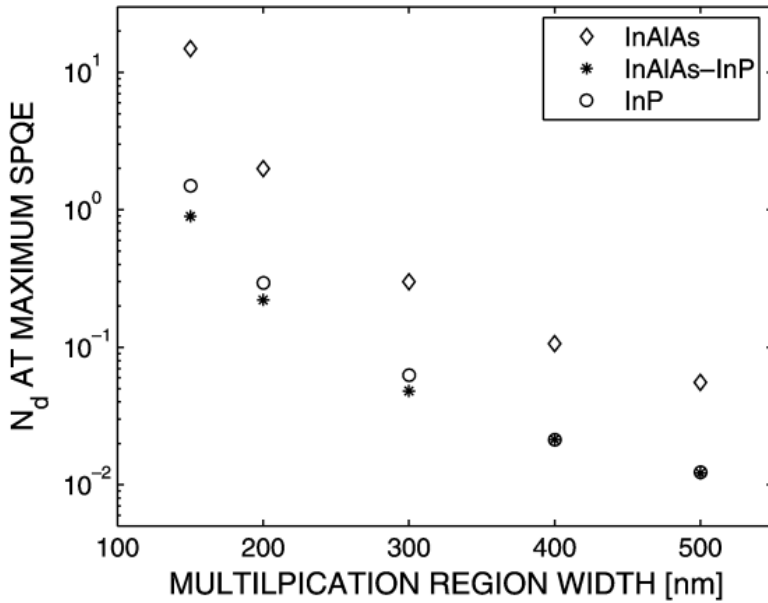


Fig. 4. Average number of dark carriers for InP, $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, and $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ –InP multiplication regions as a function of the width of the multiplication region. The values correspond to the peak SPQE.

As we indicated earlier, the main reason for the enhancement in the SPQE characteristics, as the width is increased, is the reduction in the breakdown electric field, which, in turn, leads to a reduction in dark carriers. Fig. 4 shows the average number of dark-carriers as functions of the multiplication-region width. A similar trend is seen in the case of InP multiplication regions whereby increasing the multiplication-region width is accompanied by an increase in SPQE (results are not shown). To date, a parametric model for the tunneling current in InP is not available, but at high-electric fields it is expected that the tunneling-current parameters are approximately equal for InP and InAlAs. Thus, we have approximated the dark-carrier generation in InP by using the model developed for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ [28]. However, as InP has a narrower bandgap compared to $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, the actual dark-carrier generation in InP would be higher than that predicted here and the SPQE achieved by the InP multiplication region might be smaller compared to our predictions. Nonetheless, at the high operational fields considered in our calculations, the aforementioned approximation is expected to be reasonably accurate.

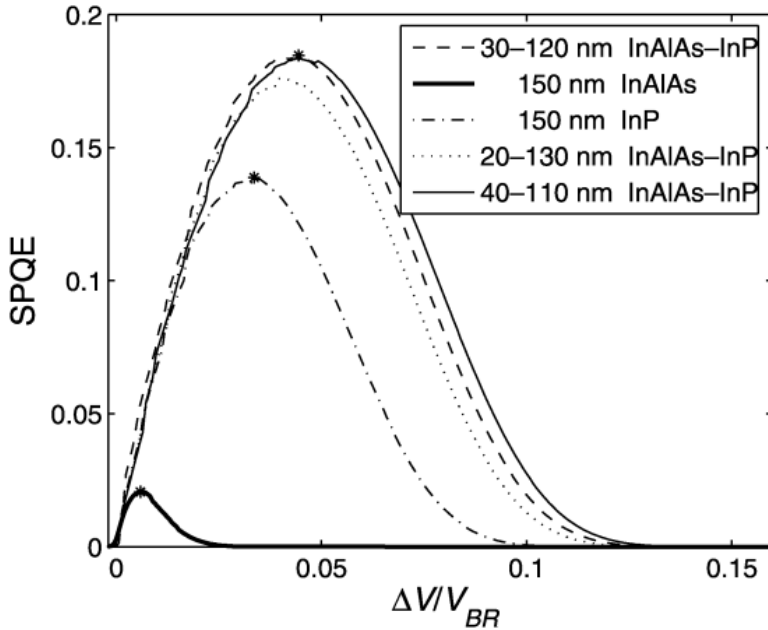


Fig. 5. SPQE for 150-nm InP (dashed-dotted curve) and 150-nm (thick solid curve) $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ multiplication regions and a 30-nm–120-nm $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ –InP heterojunction multiplication region (dashed curve). The SPQE for other combinations of the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ –InP region are also shown establishing that the 30-nm–120-nm combination is the optimal choice.

Table 1. Improvement in the Peak SPQE, and the SPQE FWHM Obtained for InP Multiplication Regions Relative to That for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ Multiplication Regions of the Same Width

	SPQE Improvement (%)	
Multiplication width (nm)	Peak	FWHM
150	573.8	294.6
200	158.8	133.7
300	37.4	79.6
400	16.1	61.5
500	8.4	54.9

Comparing an $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ multiplication region with its InP counterpart indicates that for the same width, InP achieves higher SPQE than $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, as shown by the lower two curves in Fig. 5 in the case of a 150-nm-wide multiplication region. This improvement is more pronounced in thin multiplication regions, as shown in Table I. The higher SPQE for InP is attributed to the lower breakdown electric field required by InP, and hence reduced number of dark carriers, in comparison to that required by a $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ multiplication region of the same width, as shown in Fig. 4. Additionally, in comparison to $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, InP also offers a wider SPQE versus excess-voltage curve, measured by the full-width half-maximum (FWHM), than that of $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, as shown in Table I. A large FWHM implies less variability, as a function of variation in the applied voltage, in the SPQE about its peak. Thus, in comparison to $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, InP offers less sensitivity to voltage variation.

SECTION V. Performance of SPADs With $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ --InP Heterojunction Multiplication Regions

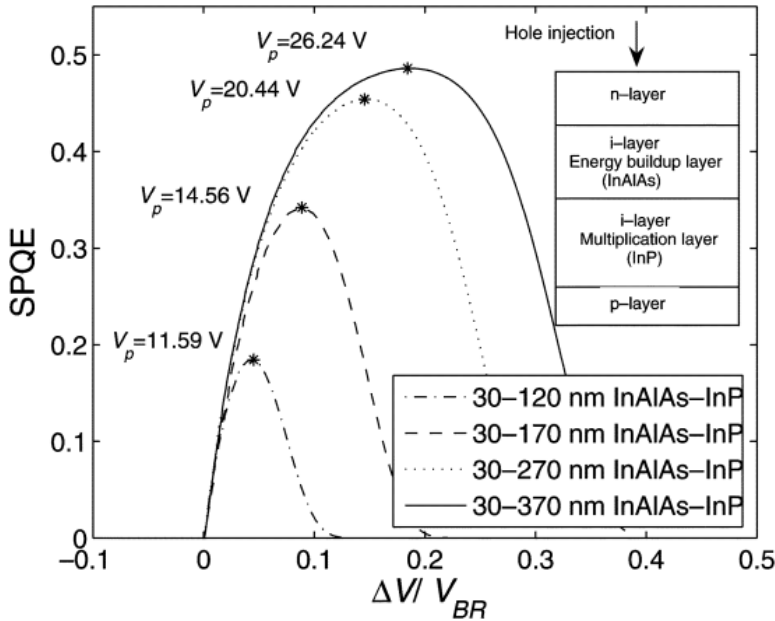


Fig. 6. SPQE as a function of the normalized excess voltage for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ --InP heterojunction multiplication regions of various widths. The reverse-bias voltage corresponding to the peak SPQE is also shown for each width. The inset shows the structure of the multiplication region. Photogenerated holes are injected into the buildup layer.

The heterojunction multiplication-region structure studied is a SAM APD. It consists of a high-bandgap $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ layer, called the *energy-buildup layer*, and a low-bandgap InP layer, called the *multiplication layer*. Consider a SAM SPAD with this heterojunction multiplication region and an InGaAs absorber. A photogenerated hole in the InGaAs layer drifts, under the effect of a weak electric field in the absorber, to the high-field $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ --InP multiplication region and enters the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ layer. As the hole is transported within the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ layer, it gradually builds up kinetic energy from the field. Once the acquired energy reaches the hole ionization threshold energy in $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, it may impact ionize according to an exponential probability law. The minimal distance, depending on the field, that the hole must travel before reaching the ionization threshold is called the *hole dead space* [32], [33].

Clearly, for a given value of the electric field, if the width of the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ layer is selected slightly less than the hole dead space, then the probability that the hole impact ionizes in the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ layer is zero. Moreover, the hole would enter the InP layer with an energy in excess of its ionization threshold energy in InP, since the ionization threshold for InP is lower than that for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$. Consequently, the hole entering the InP layer will require zero dead space and it may impact ionize *immediately* according to an exponential probability law. We emphasize that in the absence of the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ energy-buildup layer, as in a conventional multiplication region consisting solely of an InP layer, holes are injected into the InP multiplication layer “cold” (i.e., with negligible energy) and they must travel the dead space before being able to ionize. Heterojunction multiplication regions with this type of property are referred to as impact-ionization-engineered (I^2E) multiplication regions [22], [34] [35][36] [37]; they have been shown to offer significant reduction in the excess noise factor in comparison to multiplication regions that consist solely one material [34], [36]. In addition, they have been shown theoretically to exhibit enhanced breakdown probability characteristics [23], [27]. However, their performance advantage in Geiger-mode operation has not been fully established to date. We next show that heterojunction multiplication regions can offer a performance advantage in SPADs.

A. Predicted Performance Advantage

The breakdown probabilities for the multiplication region shown in the inset of Fig. 6 are calculated using the DSMT technique described in Section III-A. In this case, the ionization path densities $h_e(\xi|x)$ and $h_h(\xi|x)$ assume their respective values in each layer and the electron and hole dead-space profiles are calculated for the entire heterojunction multiplication region according to the method described [34].

The curves in Fig. 5 depict the SPQEs for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, InP and $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ –InP multiplication regions; the width is 150 nm in all cases. It is concluded that for thin multiplication regions, the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ –InP heterojunction outperforms both $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ and InP multiplication regions. However, as the width of the multiplication-region increases, the performance advantage of the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ –InP multiplication region becomes less significant and eventually approaches that of the InP multiplication region. This behavior can be substantiated and understood by examining Fig. 4, which shows that for a thick multiplication region (>400 nm), the mean number of dark carriers for the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ –InP multiplication region are almost the same as the case of InP.

As in the case of linear-mode operation and the excess noise reduction therein [36], the fractional width of the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ energy-buildup layer plays a key role in the performance in Geiger-mode operation of the device. In particular, for each overall multiplication-region width, the portion of the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ buildup layer can be chosen to maximize the SPQE. The optimization of a 150-nm multiplication region is illustrated in Fig. 5. The best SPQE performance is achieved for the choice of a 30-nm $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ buildup layer and a 120-nm InP multiplication layer. Notably, the 30-nm is approximately equal to the hole dead space of InP. This is because the optimal width of the InAlAs layer should be the distance necessary for the holes to acquire the ionization-threshold energy of InP. Consequently, holes would enter the InP layer “hot” and may ionize immediately without requiring a dead space. In the optimized structure, therefore, the hole is injected “hot” (with zero dead space) in the InP layer, thereby enhancing the probability of ionizing near the heterojunction, *viz.*, increasing the

likelihood that the first ionization occurs early on in the InP layer. Now, the fact that a hole is replaced by two holes early on enhances the breakdown probability as each hole can independently trigger a breakdown. We would therefore expect the required overall breakdown field for achieving a certain breakdown probability to be lower in an optimized heterojunction than that for the case when the first hole ionization is delayed, as in the case when the hole is injected “cold.”

To summarize, the breakdown field is lower in an optimized heterojunction multiplication region compared to a nonoptimized combination for a fixed total multiplication-region width and fixed breakdown probability.

In general, for a given width of the heterojunction multiplication region, a simple way to find the optimal width of the energy-buildup layer is as follows.

Table 2. Improvement in the Peak SPQE, and the SPQE FWHM Obtained for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ -- InP Multiplication Regions, Relative to That for InP Multiplication Regions of the Same Width

	SPQE Improvement (%)	
Multiplication width (nm)	Peak	FWHM
150	33.1	28.42
200	8.1	12.2
300	0.98	4.31
400	0.26	2.12
500	0.06	0.95

We calculate the hole dead space in InP and $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ materials, both at the breakdown field, and use them as a guide for the range of values for the choice of the optimal width of the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ energy-buildup layer. Actually, the optimal width is found to be slightly less than the hole dead space of $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, calculated at the breakdown field. For example, for the widths of 150, 200, and 300 nm, the hole dead spaces are found to be 29.9, 30.2, and 32.1 nm, respectively. On the other hand, our peak-SPQE-based optimization of the same family of structures yields energy-buildup widths of 30, 30, and 32 nm, which approximate the hole dead space at peak SPQE for the various widths considered.

Table II shows the SPQE performance advantage of the optimized $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ -- InP multiplication region compared to that of an InP multiplication region of the same width. Note that the performance advantage is present only in thin structures and then it vanishes as the width increases. This is essentially due to the diminishing role played by the dead space in thick multiplication regions. Additionally, as the operating electric field is far lower in thick multiplication layers, the impact of further reducing the field on lowering the tunneling current, which is due to the presence of the energy-buildup layer, becomes negligible in larger multiplication-layer widths. This behavior is consistent with the trend in the performance advantage of (I^2E) APDs in the context of reducing the excess noise factor [36] where the beneficial effect of the energy-buildup layer is noticeable mostly in thin structures.

We end the discussion by reiterating that the dark-carrier model [28] used in this paper takes into account the dark carriers resulting from band-to-band tunneling only, which was done as a convenient

approximation. More accurate predictions of the SPQE require the inclusion of generation-recombination dark carriers in the multiplication region as well as the absorber. However, the effect of this approximation is not expected to adversely impact the validity of the trends predicted by our model. Interface effects are also important in heterojunction multiplication regions. The valence-band discontinuity in the interface causes a diminution in the hole transition rate from the buildup layer to the multiplication layer due to an increment in hole trapping. This effect would lead to afterpulsing. To reduce this effect, fabrication efforts have been made to grade the valence-band discontinuity by inserting intermediate bandgap materials in the interface. In addition, fabrication processes are very important in reducing impurities and defects that lead to a high concentration of trapping centers. For instance, process cleanliness is critical in the fabrication of good performance detectors, since even the smallest of dust particles can lead to a poor detector performance.

SECTION VI. Conclusion

We have developed a rigorous model for calculating the single-photon quantum efficiency, SPQE, of SPADs operated in nanosecond-gated mode with arbitrary multiplication regions. The SPADs are assumed to have either a single-layer multiplication region, as the case in conventional APDs, or a two-layer, heterojunction multiplication regions, as in (I^2E) APDs. The results show that wider multiplication regions offer an improved SPQE at the expense of higher operational voltage. This trend is present in both single-layer (InP or $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$) and heterojunction ($\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ -- InP) multiplication regions. The proposition that reducing the thickness of the multiplication region in SPADs yields inferior performance was suggested earlier by this group and others [23], [24], [31]. In particular, it has already been shown that the steepness in the zero-to-one transition of the breakdown-probability versus excess-voltage curve can be improved by increasing the width the multiplication layer. However, the SPQE is not only affected by the breakdown characteristics (*viz.*, the aforementioned steepness effect) but also by the dark-count rate. The SPQE analysis presented in this paper does capture both of these factors and thus definitively establishes the benefit offered by SPADs with thicker multiplication regions.

The predicted effect of reducing the width of the multiplication region on the SPQE is primarily attributable to the increase in the breakdown electric-field in thinner multiplication regions. Indeed, the increase in the breakdown field causes the field-sensitive dark-carrier population to increase, which ultimately degrades the SPQE. Notably, the behavior of the SPQE, as a function of the multiplication-region width, is at odds with the corresponding trend in the excess-noise factor. The latter is known to decrease as the thickness of the multiplication region is reduced.

Our results show that for SPADs with $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ -- InP heterojunction multiplication regions, the width of the wide-bandgap $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ energy-buildup layer can be optimized, for a fixed total width of the multiplication region, to maximize the SPQE. In agreement with the results obtained earlier for optimizing (I^2E) multiplication regions for minimal excess-noise factor, the optimal width of the buildup layer is approximately equal to the injected-carrier's dead space. For example, in a 200-nm multiplication region, the optimal buildup-layer width that maximizes the SPQE is 30 nm and the hole dead space is 30.2 nm. Compared with a single-layer InP multiplication region, the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ -- InP heterojunction achieves tangible improvement in the SPQE in thin multiplication regions (<200 nm). In the thicker multiplication regions, on the other hand, this improvement is

negligible. This behavior can be understood in the context of increased significance of the dead-space effect in thin multiplication regions.

The effect of afterpulsing was neglected in our model primarily due to the lack of precise knowledge of the dependence of the carrier-release time on the field. As afterpulsing generally increases as the electric field increases [38], the inclusion of afterpulsing in our model would not change our conclusions for single-layer multiplication regions. However, in the case of $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ -- InP heterojunction, we expect that hole trapping at the heterojunction may accentuate afterpulsing due to the presence of the $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ -- InP potential barrier as well as high concentration of trapping levels at the heterojunction [8]. However, whenever the information of the dependence of the carrier-release time on the field becomes available, it is straightforward to include afterpulsing in our analysis by merely modifying the parameter N_d to include released carriers. This can be done following the approach outlined by Kang et al. [25], which assumes knowledge of the carrier-release rate. A full afterpulsing analysis will be pursued in the future.

References

1. C. P. Morath, K. Vaccaro, W. R. Clark, W. A. Teynor, M. A. Roland, W. Bailey, "Performance characterization of an InGaAsInP single photon avalanche diode", *Proc. 49th Annu. Meeting SPIE: Applications of Digital Image Processing XXVII*, pp. 100-111, 2004.
2. A. Biswas, W. H. Farr, "Detectors for ground-based reception of laser communications from mars", *Proc. 17th Annu. Meeting IEEE Lasers and Electro-Optics Society (LEOS)*, vol. 1, pp. 74-75, 2004.
3. A. Spinelli, L. M. Davis, H. Dautet, "Actively quenched single-photon avalanche diode for high repetition rate time-gated photon counting", *Rev. Sci. Instrum.*, vol. 67, pp. 55-61, 1996.
4. D. Stucki, G. Ribordy, H. Z. A. Stefanov, J. G. Rarity, "Photon counting for quantum key distribution with peltier cooled InGaAsInP APDs", *J. Mod. Opt.*, vol. 48, pp. 1967-1981, 2001.
5. M. B. Nasr, B. E. A. Saleh, A. V. Sergienko, M. C. Teich, "Dispersion-cancelled and dispersion-sensitive quantum optical coherence tomography", *Opt. Exp.*, vol. 12, pp. 1353-1362, 2004.
6. H. Zbinden, H. Bechmann-Pasquanucci, N. Gisin, G. Ribordy, "Quantum cryptography", *Appl. Phys. B*, vol. 67, pp. 743-748, 1998.
7. J. M. Merolla, Y. Mazurenko, J. P. Goedgebuer, L. Durafourg, H. Porte, W. T. Rhodes, "Quantum cryptographic device using single-photon phase modulation", *Phys. Rev. A*, vol. 60, pp. 1899-1905, 1999.
8. S. Cova, M. Ghioni, A. Lotito, I. Rech, F. Zappa, "Evolution and prospects for single-photon avalanche diodes and quenching circuits", *J. Mod. Opt.*, vol. 51, pp. 1267-1288, 2004.
9. G. Ribordy, N. Gisin, O. Guinnard, D. Stucki, M. Wegmuller, H. Zbinden, "Photon counting at telecom wavelengths with commercial InGaAsInP avalanche photodiodes: Current performance", *J. Mod. Opt.*, vol. 51, pp. 1381-1398, 2004.
10. *Single Photon Counting Module Datasheet*, [online] Available: .
11. M. Ghioni, A. Lacaita, G. Ripamonte, S. Cova, "All-silicon avalanche photodiode sensitive at $1.3\ \mu\text{m}$ with picosecond time resolution", *IEEE J. Quantum Electron.*, vol. 28, no. 12, pp. 2678-2681, Dec. 1992.
12. A. Lacaita, F. Zappa, S. Cova, P. Lovati, "Single-photon detection beyond $1\ \mu\text{m}$: Performance of commercially available InGaAsInP detectors", *Appl. Opt.*, vol. 35, pp. 2986-2996, 1996.

13. I. Prochazka, K. Hamal, B. Sopko, "Recent achievements in single photon detectors and their applications", *J. Mod. Opt.*, vol. 51, pp. 1289-1313, 2004.
14. P. A. Hiskett, G. S. Buller, A. Y. Loudon, J. M. Smith, I. Gontijo, A. C. Walker, P. D. Townsend, M. J. Robert, "Performance and design of InGaAsInP photodiodes for single-photon counting at 1.55 μm ", *Appl. Opt.*, vol. 39, no. 36, pp. 6818-6829, 2000.
15. A. Y. Loudon, P. A. Hiskett, G. S. Buller, R. T. Carline, D. C. Herbert, W. Y. Leong, J. G. Rarity, "Enhancement of the infrared detection efficiency of silicon photon-counting avalanche photodiodes by use of silicon germanium absorbing layers", *Opt. Lett.*, vol. 27, pp. 219-221, 2002.
16. M. M. Hayat, W. L. Sargeant, B. E. A. Saleh, "Effect of dead space on gain and noise in Si and GaAs avalanche photodiodes", *IEEE J. Quantum Electron.*, vol. 28, no. 5, pp. 1360-1365, May 1992.
17. M. A. Saleh, M. M. Hayat, B. E. A. Saleh, M. C. Teich, "Dead-space-based theory correctly predicts excess noise factor for thin GaAs and AlGaAs avalanche photodiodes", *IEEE Trans. Electron Devices*, vol. 47, no. 3, pp. 625-633, Mar. 2000.
18. P. Yuan, C. C. Hansing, K. A. Anselm, C. V. Lenox, H. Nie, A. L. Holmes Jr., B. G. Streetman, J. C. Campbell, "Impact-ionization characteristics of IIIV semiconductors for a wide range of multiplication region thicknesses", *IEEE J. Quantum Electron.*, vol. 36, no. 2, pp. 198-204, Feb. 2000.
19. K. F. Li, D. S. Ong, J. P. R. David, G. J. Rees, R. C. Tozer, P. N. Robson, R. Grey, "Avalanche multiplication noise characteristics in thin GaAs p-i-n diodes", *IEEE Trans. Electron Devices*, vol. 45, no. 10, pp. 2102-2107, Oct. 1998.
20. M. Ma, S. Wang, X. Li, A. Anselm, X. G. Zheng, A. L. Holmes Jr., J. C. Campbell, "Monte-Carlo simulation of low-noise avalanche photodiodes with heterojunctions", *J. Appl. Phys.*, vol. 92, pp. 4791-4795, 2002.
21. C. Groves, J. P. R. David, G. J. Rees, D. S. Ong, "Modeling of avalanche multiplication and noise in heterojunction avalanche photodiodes", *J. Appl. Phys.*, vol. 95, pp. 6245-6251, 2004.
22. N. Li, R. Sidhu, X. Li, F. Ma, X. Zheng, S. Wang, G. Karve, S. Demiguel, A. L. Holmes Jr., J. C. Campbell, "InGaAsInAlAs avalanche photodiode with undepleted absorber", *Appl. Phys. Lett.*, vol. 82, pp. 2175-2177, 2003.
23. O.-H. Kwon, M. M. Hayat, J. C. Campbell, B. E. A. Saleh, M. C. Teich, "Optimized breakdown probabilities in $\text{Al}_{0.6}\text{Ga}_{0.4}\text{As}/\text{GaAs}$ heterostructure avalanche photodiodes", *IEEE Electron Device Lett.*, vol. 25, pp. 599-601, 2004.
24. S. Wang, F. Ma, X. Li, G. Karve, X. Zheng, J. C. Campbell, "Analysis of breakdown probabilities in avalanche photodiodes using a history-dependent analytical model", *Appl. Phys. Lett.*, vol. 82, pp. 1971-1973, 2003.
25. Y. Kang, H. X. Lu, Y. H. Lo, D. S. Bethune, W. P. Risk, "Dark count probability and quantum efficiency of avalanche photodiodes for single-photon detection", *Appl. Phys. Lett.*, vol. 83, pp. 2955-2957, 2003.
26. N. Z. Hakim, B. E. A. Saleh, M. C. Teich, "Generalized excess noise factor for avalanche photodiodes of arbitrary structure", *IEEE Trans. Electron Devices*, vol. 37, pp. 599-610, 1990.
27. M. M. Hayat, Ü Sakoğlu, O.-H. Kwon, S. Wang, J. C. Campbell, B. E. A. Saleh, M. C. Teich, "Breakdown probabilities for thin heterostructure avalanche photodiodes", *IEEE J. Quantum Electron.*, vol. 39, pp. 179-185, 2003.
28. G. Karve, S. Wang, F. Ma, X. Li, J. C. Campbell, "Origin of dark counts in $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ avalanche photodiodes operated in geiger mode", *Appl. Phys. Lett.*, vol. 86, pp. 631-633, 2005.

29. K. A. McIntosh, J. P. Donnelly, D. C. Oakley, A. Napoleone, S. D. Calawa, L. J. Mahoney, K. M. Molvar, E. K. Duerr, S. H. Groves, D. C. Shaver, "InGaAsP/InP avalanche photodiodes for photon counting at $1.06\ \mu\text{m}$ ", *Appl. Phys. Lett.*, vol. 81, pp. 2505-2507, 2002.
30. G. Karve, X. Zheng, X. Zhang, X. Li, S. Wang, F. Ma, A. Holmes, J. C. Campbell, G. S. Kinsey, J. C. Boisvert, T. D. Isshiki, R. Sudharsanan, D. S. Bethune, W. P. Risk, "Geiger mode of an $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ avalanche photodiodes", *IEEE J. Quantum Electron.*, vol. 39, no. 10, pp. 1281-1286, Oct. 2003.
31. J. S. Ng, C. H. Tan, J. P. David, G. J. Rees, "Theoretical study of breakdown probability in single photon avalanche diode", *Proc. 16th Annu. Meeting IEEE Lasers and Electro-Optics Soc.*, vol. 2, pp. 773-774, 2003.
32. Y. Okuto, C. R. Crowell, "Ionization coefficients in semiconductors: A nonlocalized property", *Phys. Rev. B*, vol. 10, pp. 4284-4296, 1973.
33. B. E. A. Saleh, M. M. Hayat, M. C. Teich, "Effect of dead space on the excess noise factor and time response of avalanche photodiodes", *IEEE Trans. Electron Devices*, vol. 37, no. 9, pp. 1976-1984, Sep. 1990.
34. M. M. Hayat, O.-H. Kwon, S. Wang, J. C. Campbell, B. E. A. Saleh, M. C. Teich, "Boundary effects on multiplication noise in thin heterostructure avalanche photodiodes: Theory and experiment", *IEEE Trans. Electron Devices*, vol. 49, no. 12, pp. 2114-2123, Dec. 2002.
35. M. Wang, R. Sidhu, X. G. Zheng, X. Li, X. Sun, A. L. Homes Jr., J. C. Campbell, "Low-noise avalanche photodiodes with graded impact-ionization-engineered multiplication region", *IEEE Photon. Technol. Lett.*, vol. 13, no. 12, pp. 1346-1348, Dec. 2001.
36. M. Kwon, M. M. Hayat, S. Wang, J. C. Campbell, A. Holmes, B. E. A. Saleh, M. C. Teich, "Optimal excess-noise reduction in thin heterojunction $\text{Al}_{0.6}\text{Ga}_{0.4}\text{As}/\text{GaAs}$ avalanche photodiodes", *J. Quantum Electron.*, vol. 39, no. 10, pp. 1287-1296, Oct. 2003.
37. C. Groves, J. P. R. David, G. J. Rees, D. S. Ong, "Modeling of avalanche multiplication and noise in heterojunction avalanche photodiodes", *J. Appl. Phys.*, vol. 95, pp. 6245-6251, 2002.
38. S. Cova, M. Ghioni, A. Lacaita, C. Samori, F. Zappa, "Avalanche photodiodes and quenching circuits for single-photon detection", *Appl. Opt.*, vol. 35, pp. 1956-1976, 1996.